CLAIMS

[Claim(s)]

[Claim 1] It is made the storage system which has the dispersed system component containing a host processor. Said host processor runs the software application which generates renewal of a record. When it has a data mover, and was combined with the 1st storage controller and an error situation arises in said storage system, In the storage system which performs an approach effective in a machine in order that said data mover may adjust said dispersed problem decision between system components (a) The step which produces the I/O operation for the renewal of a record generated by said software application, (b) The step which memorizes said renewal of a record in said 1st storage controller according to said produced I/O operation, (c) The step which maintains the control information relevant to said renewal of a record in said 1st storage controller, (d) The step which reads into said data mover said control information related [which was related and record-updated] from said 1st storage controller in preparation for copying said renewal of a record remotely, (e) The step which detects the storage system error situation in said data mover, (f) The step which generates diagnostic-state storage channel command word (CCW) for said host processor and said 1st storage controller from said data mover, (g) Said fault information in each system component is an approach characterized by correlating according to the detected error situation including the step which catches fault information in said host processor and said storage controller.

[Claim 2] Said step (f) is an approach according to claim 1 characterized by including the step which freezes system environment until said step (g) is completed.

[Claim 3] Said approach is an approach according to claim 2 characterized by including the step which analyzes fault information further in order to determine the cause of the (h) error.
[Claim 4] Said step (e) is an approach according to claim 2 characterized by including the step which receives the error code showing an error type in said data mover.
[Claim 5] Said step (g) is an approach according to claim 2 characterized by including the step which memorizes host-processor fault information in the circulation trace buffer in said data mover.
[Claim 6] Said step (g) is an approach according to claim 2 characterized by including the step which memorizes the 1st storage controller fault information in the circulation trace buffer in said 1st storage controller.
[Claim 7] Said fault information is an approach according to claim 2 characterized by including the hardware, the software, and the microcode control structure for giving debugging.
[Claim 8] Said step (f) is an approach according to claim 2 characterized by including the step which generates said diagnostic-state storage CCW for the 2nd storage controller in a remote location.
[Claim 9] It is made the storage in which the computer read for memorizing data mover application in order to answer a storage system error and to combine a condition storage diagnosis ranging over many system components is possible. In the storage containing the host processor combined with the storage controller for said system component to manage the renewal of a record generated by the software application to a direct access storage device (DASD) The means for producing the I/O operation to the renewal of a record generated in said software application, The means for memorizing said renewal of a record for said storage controller according to said produced I/O operation, The means for maintaining the control information relevant to said

renewal of a record in said storage controller, The means for reading into said data mover said control information related [which was related and record-updated] in preparation for copying said renewal of a record remotely, The detection means for detecting a storage system error situation and communicating said error situation to said data mover, The condition storage means for generating diagnostic-state storage channel command word (CCW) for said host processor and said storage controller from said data mover, The prehension means for catching fault information in said host processor and said storage controller, An implication and said fault information in each system component of the system components of said large number are data mover application characterized by correlating according to the detected error situation.

[Claim 10] Said condition storage means is data mover application according to claim 9 characterized by freezing the software application relevant to said error situation until the operation and said prehension means of said storage controller complete prehension of fault information.

[Claim 11] Said detection means is data mover application according to claim 9 characterized by receiving the error code showing the error type of said error situation in said data mover.

[Claim 12] Said prehension means is data mover application according to claim 9 characterized by memorizing host-processor fault information in the circulation trace buffer in said data mover.

[Claim 13] Said prehension means is data mover application according to claim 12 characterized by memorizing storage controller fault information in the circulation trace buffer in said storage controller.

[Claim 14] It is made the data storage system for adjusting fault information between [relevant to the generated error situation] system components. In the data storage system containing one or more storage

controllers combined with the nonvolatile storage for said system component to memorize renewal of a record The host processor which runs a software application is included. Said software application is what generates renewal of a record and sends the I/O operation for the usual storage in said nonvolatile storage to said one or more storage controllers. Said host processor reads said renewal of a record in said one or more storage controllers. It is a thing containing the data mover for assembling said renewal of a record in a group, in order to send to the remote storage system for recovery at the time of disaster. Said data mover is what receives the error code showing the type of the produced error situation from one of said the system components. It is what generates the condition storage command for making said related system component interrupt processing of renewal of a record temporarily in order that said data mover may collect fault information in said related system component. And that said fault information is what is correlated between system components according to a condition storage command and said data mover The trace queue for memorizing the fault information relevant to said data mover, The data storage system characterized by including the control section for managing the renewal of a record read into said data mover, and two or more buffers for memorizing the header to which said renewal of a record and they relate.

[Claim 15] The data storage system according to claim 14 characterized by including the remote data mover combined so that the group of renewal of a record might be received from said data mover.

[Claim 16] Said nonvolatile storage is a data storage system according to claim 14 characterized by being a direct access storage device (DASD).

[Claim 17] It is the data storage system according to claim 15

characterized by answering said condition storage command when an error situation produces said remote data storage system in said remote data storage system including the remote data storage system combined so that the group of renewal of a record might be received from said remote data mover for storage.

[Claim 18] Said remote data storage system is a data storage system according to claim 17 characterized by including a remote host processor, the remote storage controller combined with said remote host processor, and the remote nonvolatile storage for being combined with said remote storage controller and memorizing the group of said renewal of a record.

[Claim 19] Said fault information is software, hardware, and a data storage system according to claim 14 characterized by including a microcode control structure.

[Claim 20] The secondary site which receives a copy of data for recovery at the time of the primary site for processing data, and disaster, It has software, hardware, and the distributed-system component that had the microcode control structure incorporated. It is made a recovery data storage system at the time of the disaster which gives error debugging by carrying out the trigger of adjustment and prehension of the fault information relevant to the generated specific error situation distributed-system component mutual [said]. Two or more storage controllers by which each has cache memory, a control buffer, and a trace buffer, It is what sends the I/O operation to each renewal of a record to said two or more storage controllers in order that said application may generate renewal of a record and may write said renewal of a record in two or more storage controllers including the host processor which runs application. Said host processor is a thing containing the primary data mover for reading said renewal of a record in

said two or more storage controllers, and assembling it in the group of a self-description record set. Said primary data mover is what receives the error code showing the type of the produced error situation from one of said the system components. In order that said primary data mover may collect fault information It is what generates the diagnostic-state storage channel command word for making said related system component interrupt processing of renewal of a record temporarily towards the distributed-system component relevant to said error situation. Said fault information is what is correlated according to diagnostic-state storage channel command word, Two or more primary direct access storage devices combined with said two or more primary storage controllers (DASD), The secondary-data mover which is connected so that the group of a self-description record set may be received, and answers said diagnostic-state storage channel command word, The secondary host processor which is memorized by said secondary-data mover and answers said diagnostic-state storage channel command word, Two or more secondary-storage controllers which are combined with said host processor and answer said diagnostic-state storage channel command word, Secondary [for memorizing the group of said self-description record set / two or more / DASD] is included. Said primary data mover The trace queue for memorizing the fault information relevant to said primary data mover, It is a recovery data storage system at the time of the disaster characterized by including the control section for managing the renewal of a record read into said primary data mover, and two or more buffers for memorizing the header to which said renewal of a record and they relate.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the improved condition storage operation over the component of a large number in a hierarchy data storage system, if it says generally and will say in detail about hierarchy data storage.

[0002]

[Description of the Prior Art] Generally, data processing system is required to memorize a lot of data (or record) in relation to processing of data. In addition, those data are efficiently accessible, and can be corrected, and must be able to be restored. Data storage is divided into the level from which some differ in order [being effective] to reach and to perform cost effectiveness- data storage, i.e., a hierarchy target. The data storage of the 1st level, i.e., a record level, means random access memory (DRAM or SRAM) dynamic [electronic memory and usual] or static. Electronic memory takes the format of a semiconductor integrated circuit. Then, it can be remembered that millions of bytes of data are even for access for several nanoseconds to such a data byte in

each circuit. Since access is electronic on the whole, the electronic memory performs quickest access.

[0003] The data storage of the 2nd level usually means a direct access storage device (DASD). It consists of an optical disk, and since "1" which forms the bit of data, or "0" is expressed, DASD memorizes a data bit on a disk side as magnetic or a spot of micrometer (micron) size changed magnetically or optically, for example. One or more disks covered as MAG DASD is also for remnant magnetism material are included. It is equipped with those disks pivotable in the protected environment. Each disk is divided into many concentric circle trucks, i.e., the circle which approached. Data are memorized in serial for every bit along each truck. the time of generally a disk rotating the access mechanism known as a head disk assembly (HDA) through its read/write head including one or more read/write heads -- those trucks -- crossing -- moving -- the front face of a disk -- and in order to transmit data from the front face, it is prepared in each DASD. Generally DASD can remember several giga cutting tool's data that access (later in the magnitude of several figures than electronic memory) of several mm second to such data is also. Since it is necessary to position physically access to the data memorized by DASD in a disk and the data storage location of a request of HDA, it is a low speed.

[0004] The data storage of the 3rd level, i.e., a low, includes a tape library or a tape, and a DASD library. Since a robot or an operator needs to choose and load the requested data carrier, access to data is slow much in a library. The advantage is reduction of the cost to several terabytes of a very large data memory storage function, for example, data memory storage function. Generally a tape storage is used for the purpose of backup. That is, the data memorized on a hierarchy's 2nd level are copied on a magnetic tape for protection. a tape top -- or

access to the data memorized to the library is the thing of a second unit in current.

[0005] Since loss of data may become fatal for a company, it is indispensable for many companies to have the data copy of backup. Time amount required to recover the data lost on primary storage level is also an important recovery consideration matter. The dual copy which is another format of backup gives amelioration of a rate in a tape or library backup. An example of a dual copy contains what additional DASD with which data are written in is formed for (called mirroring). Then, when primary [DASD] breaks down, it will depend for data on secondary [DASD]. The fault of this approach is that the number of required DASDs becomes two times.

[0006] It includes that other data backup substitution approaches of conquering the need of making a store into two times write data in the redundancy array (RAID) configuration of a low-price disk. In this example, data are written in so that it may be distributed among many DASDs. When one DASD breaks down, the lost data can be recovered by using remaining data and a remaining error correction procedure. The RAID configuration from which some differ is usable current.

[0007] Furthermore, the backup solution which performs large-scale protection is a remote dual copy which needs to carry out the shadow of the primary data memorized by primary [DASD] in a secondary location or a remote location. The distance which divides a primary location and a secondary location can be changed from the thing over a fire wall to several km about synchronous data communication depending on the level of the risk which can be received for a user. In addition to performing a backup data copy, a secondary location or a remote location must also have sufficient system information to succeed the processing to the first order system, when the disable of the first order

system is carried out. Since a single storage controller does not write data in both the primary DASD string in a primary site and a secondary site, and a secondary DASD string, this is partially natural. Instead, primary data are memorized on the primary DASD string connected to the primary storage controller, and, on the other hand, a secondary data is memorized on the secondary DASD string connected to the secondary-storage controller.

[0008] A remote dual copy belongs to either two general categories, i.e., a synchronous thing, or an asynchronous thing. Before a synchronous remote copy ends sending primary data to a secondary location, and primary DASD input / output (I/O) operation (channel-end (CE) / unit end (DE) is given to a primary host), it makes it possible to check receipt of such data. Therefore, a synchronous remote copy makes the I/O response time primary [DASD] slow, while waiting for a secondary check. Primary I/O answering delay increases in proportion to the distance between a first order system and a second-order system (it is an element which restricts the remote distance to about tenkm). However, a synchronous remote copy supplies the adjusted data [in / that comparatively few system overheads are also / a secondary site] one by one.

[0009] An asynchronous remote copy is that (channel-end (CE) / unit end (DE) is supplied to a primary host) which I/O operation primary [DASD] completes, before data are checked in a secondary site, and it gives good primary application system performance. Therefore, it could depend for the primary DASDI/O response time on the distance to a secondary site, but may be [the secondary site] thousands of km away from the primary site. However, since the data which won popularity to the secondary site have often reached in different sequence from the data written on primary [DASD], in order to guarantee the adjustment

of a data sequence, a lot of system overheads are required. The data which were under conveyance between the primary location and the secondary location may be lost as a result of the failure in a primary site.

[0010] A recovery method includes the remote dual copy which data do not come to accept remotely and by which they are backed up also continuously at the time of the data disaster introduced recently. In a common remote dual copy system, the primary processor of a large number connected to many primary storage controllers by many serials or the parallel communication link exists. The each first storage controller has the string primary [DASD] connected to it. The same processing system may exist in a remote secondary site.

[0011] When there is an increment in the potential distance to which the complexity of the system introduced by the remote copy increases and relates, debugging of the problem of hardware, a microcode, or software becomes very complicated. The usual debugging technique carries out time amount consumption extremely as signs of a logical error in the field of the hardware in system excess time amount which such a problem usually produced after enough, software, or a microcode. Furthermore, system debugging is further complicated by running the software distributed by use of the channel escape box containing an extended enterprise system connection (ESCOM) channel ranging over both the first order system and the second-order system, in order to extend a communication link from hundreds of meters anywhere in the world. Therefore, the magnitude of distributed debugging becomes what is not almost manageable.

[0012] Therefore, to offer the approach and equipment for adjusting the problem decision between components of a distributed system is desired. For example, a data mover is directed to the host processor and application of the control device which produced the error, and relation,

and in order to adjust hardware, software, or the problem decision between microcodes, it makes the condition of that keep it, when returning the result in which an error has a control device.

[0013]

[Problem(s) to be Solved by the Invention] The purpose of this invention is to offer the design and approach of having been improved for adjusting the condition storage between distributed components in computer system.

[0014] Another purpose of this invention is to offer [between a host processor and storage controllers or] the condition storage to which between storage controllers was adjusted.

[0015] Further, another purpose is to generate the channel command word (CCW) for making a condition keep it in a host processor and at least one storage controller from a host processor, in order to debug an error, or in order [of this invention] to analyze system performance.

[0016]

[Means for Solving the Problem] According to the 1st example of this invention, the approach for adjusting problem generating between [in a storage system] distributed-system components is given. A storage system contains the host processor which runs host application and has a data mover. A host processor is combined with the storage controller for performing renewal of a record. The approach includes generating I/O operation in a storage controller from host application, in order to memorize the renewal of a record generated by host application. The control information containing the record header relevant to the renewal of a record is maintained by the storage controller. It equips copying renewal of a record to predetermined time amount remotely with a data mover, and it reads the control information of the relation to the group of renewal of a record, and it. When an error situation arises in the **

component of the system, the error situation is detected in a data mover according to the received error code. Answering the receipt of an error code, a data mover generates "diagnostic-state storage" channel command word (CCW) for a host processor and a storage controller. Fault information is immediately caught in a data mover and a storage controller. Then, the fault information is correlated according to the detected error situation.

[0017] Another example of this invention adjusts fault information between [relevant to the error situation which the data storage system produced in the data storage system] system components. Those system components contain software, hardware, and a microcode control structure. A system component contains one or more storage controllers combined with the host processor and the nonvolatile storage, in order to process and memorize renewal of a record. A host processor runs the application which generates renewal of a record and sends I/O operation to one or more storage controllers, and memorizes renewal of a record on a nonvolatile storage.

[0018] Furthermore, a host processor has a data mover for assembling those renewal of a record in the group of a self-description record, in order to read renewal of a record in one or more storage controllers, and in order to send to a remote storage system for the purpose of recovery at the time of disaster. When an error arises, a data mover receives the error code showing the type of the produced error situation from one of the system components. It is answered, and while collecting fault information, a data mover generates a condition storage command to the system component relevant to the error situation in order to make the system component relevant to an error situation interrupt processing of renewal of a record temporarily. The fault information over each system component is correlated according to a condition storage command. A

data mover has the trace queue which memorizes the control section for managing the renewal of a record written in the fault information and the data mover relevant to the data mover. Moreover, two or more buffers for memorizing the header of the renewal of a record and those relation are formed in a data mover.

[0019]

[Embodiment of the Invention] Common data processing system is easy to take the format of a host processor like the IBM system for running the data functional storage management subsystem (DFSMS/MVS) software of MVS for calculating and processing data / 360, or the IBM system / 370 processor. In addition, as for the processor, at least one IBM3990 storage controller is connected, and the storage controller includes one or more cache memory types built into a memory controller and it. Furthermore, a storage controller is connected to the group of a direct access storage device (DASD) like IBM3380 or 3390DASDs. Although a host processor gives the remarkable count force, a storage controller is transmitted efficiently, carries out the stage/DESUTEJI of the big database, and changes it, and it gives a function required to access on the whole.

[0020] Drawing 1 shows a recovery system 100 at the time of asynchronous disaster including the primary site 121 and a remote site, or the secondary site 131. The primary site 121 contains IBMES/9000 which run the primary processor 101, for example, DFSMS/MVS host software. Furthermore, the primary processor 101 contains application software (APPLN1, APPLN2) 102 and 103, for example, IMS, and DB2 application, and the primary data mover (PDM) 104. The common SHISUPU REXX clock 107 for performing general reference to all the applications (102 103) that run in a processor 101 is contained in the primary processor 101. All system clocks or time sources (not shown)

guarantee that processing of all time-dependents is clocked proper about mutual synchronizing with the SHISUPU REXX clock 107. The primary storage controller 105 synchronizes with the resolution suitable for guaranteeing distinction between the renewal time amount of record writing so that the time stamp value with two same continuous write-in I/O operation to the single primary storage controller 104 may not be shown. the resolution (it is not precision) of the SHISUPU REXX clock 107 is critical -- certain ** Since write-in I/O operation is not generated there, PDM104 connected to the SHISUPU REXX clock 107 does not need to synchronize with the SHISUPU REXX clock 107. When the primary processor 101 has "single only refer to the time amount" (for example, multiprocessor ES / 9000 single system), the SHISUPU REXX clock 107 is unnecessary.

[0021] Two or more primary storage controllers 105, for example, an IBM3990 model 6 storage controller, are connected to the primary processor 101 through two or more channels, for example, optical fiber channel. It is connected to each temporary storage controller 105, at least one string, for example, IBM3390DASD, primary [DASD / 106]. The primary storage controller 105 and primary [DASD / 106] form a primary storage subsystem. The each first storage controller 105 and primary [DASD / 106] do not need to be separate equipment, and it is easy to be combined with a single drawer.

[0022] For example, the secondary site 131 left and installed also thousands of km from the primary site 121 contains the secondary processor 111 which has the secondary-data mover (SDM) 114 like the primary site 121. Apart from this, primary and a secondary site may be the same locations, and the primary data mover 104 and the secondary-data mover 114 may be on a single host processor further (secondary [DASD / 116] may be in the exterior of a fire wall). Two or

more secondary-storage controllers 115 are connected to the secondary processor 111 through a channel, for example, the optical fiber channel known for the field. Secondary [two or more / DASD / 116] and control information DASD 117 are connected to the secondary-storage controller 115. The secondary-storage controller 115 and DASDs 116 and 117 constitute a secondary-storage subsystem.

[0023] The primary site 121 communicates with the secondary site 131 through a communication link 108. Furthermore, if it says in detail, the primary processor 101 will transmit data and control information to the secondary processor 111 with a communications protocol 108, for example, a virtual telecommunications access method (VTAM) communication link. A communication link 108 is realizable with many suitable correspondence procedures including a telephone (T1, T3 Rhine), wireless, wireless / telephone, microwave, a satellite, etc.

[0024] The asynchronous-data shadow system 100 collects the control data from the primary storage controller 105 so that the sequence of all data writing primary [DASD / 106] may be held and it may be applied to secondary [DASD / 116 (data write-in sequence is held over all primary storage subsystems)]. The data and control information which are sent to the secondary site 131 must be so enough that the need does not already have existence of the primary site 121 in order to hold data consistency.

[0025] Applications 102 and 103 generate data or renewal of a record. The renewal of a record is collected by the primary storage controller 105, and is read by PDM104. Each of the primary storage controller 105 gives those renewal of a record to PDM104 through a group part opium poppy and unspecified "read" request primary [DASD / 106] for each renewal of a record to an asynchronous remote data shadowing session. The amount of the data from the primary storage controller 105 to

PDM104 further transmitted between each of the each first storage controller 105 and the primary processor 101 is made into max by a transfer of renewal of a record being controlled and optimized by PDM104, and minimizing much start I/O operation. PDM104 can change the unspecified time interval between "read", in order to control the flow of the renewal of a record to secondary [this / primary storage controller / host optimization, and secondary / DASD].

[0026] Conveying those renewal of a record to collecting renewal of a record by PDM104 and SDM114 needs to convey those renewal of a record between specific time intervals to secondary [DASD / 116] through all primary storage subsystems with enough control data to reconfigurate a record write-in sequence primary [DASD / 106], maintaining the adjustment of data. Reconstruction of a record write-in sequence primary [DASD / 106] is attained by sending a self-description record to SDM114 from PDM104. SDM114 investigated the self-description record, and all the records between given time intervals have been lost, or it is determined whether it is imperfect.

[0027] Drawing 2 shows the component of the primary (or secondary) storage controller 105. As for the storage controller 105, the each has the multi-pass storage directors 262 and 268 including the dual cages 236 and 238 (in order to raise availability, an electric power supply is carried out separately). those multi-pass storage directors -- two storage pass (SP) -- it has 248, 250, and 252 and 254. Each storage pass can interpret the channel command word (CCW) received from the host processor, and can control connected DASD.

[0028] Furthermore, each cages 236 and 238 contain the shared-control array (SCA) storage 264 and 266. Each SCA 264 and 266 memorizes other SCA prepared in other cages in the storage controller 105, common status information, and local status information. Thus, when a

fault condition does not exist in both both [one side or] 236 and 238, the contents of SCA 264 and 266 are the same. The common shared data contains the equipment status and re-connection data other than the various external registers used by the microprocessor contained in each of the storage pass 248, 250, 252, and 254. SCA data are accessible for all the four storage pass 248, 250, 252, and 254.

[0029] The cache 258 and nonvolatile storage (NVS) 260 which are high-speed memory are formed in cages 236 and 238, respectively. The data which move between a host and DASD are temporarily memorizable into a cache 258, in order to improve a system throughput. In the field, the data received from the host can be copied to NVS260, for example, in order to give the DASD high-speed write-in function in which it is known well. For example, in order to raise the availability of data further, other configurations which prepare a cache and NVS memory in each cages 236 and 238 are known.

[0030] Drawing 3 sets 1 ** 248 of four storage pass in the storage controller 105 to 348, and shows it by the block graphic form formula. Four storage pass 248, 250, 252, and 254 is altogether the same, and only the storage pass 348 which is one of those [the] will be explained here. The storage pass 348 is connected to two or more DASDs by the up port 310 by 8* 2 switch 362 and the lower port 312. The data transfer between the multi-pass storage director 362 and one of the DASDs is produced through the automatic data transfer circuit (ADT) 374 supported by the rate modification buffer 376 if needed at the time of direct DASD operation. The rate modification buffer 376 compensates the difference of the rate of data transfer and the rate of channel-CPU communication by DASD. Generally, a channel transfer is much more nearly high-speed than a DASD-controller transfer, and is produced.

[0031] Moreover, it has the microprocessor 370 which functions as the

storage pass 348 controlling all the operation produced on the storage pass. The microprocessor 370 executes microinstruction loaded to control storage (not shown) through the external support device. The port adapter (PA) 372 gives the control data pass for the data transfer between a cache 358, a nonvolatile storage (NVS) 360, a microprocessor 370, the up port 310, and the lower port 312.

[0032] A cache 358 is in the cache, and includes two ports (an up port, lower port) which make possible two coincidence data transfer from the cache. For example, it is possible to transmit data to a channel 310 from a cache 358, transmitting data to a cache 358 from DASD. Generally, data transfer is succeeded without the further break in from a microprocessor 370 by the function of the automatic data transfer circuit 374 until it is set up by the microprocessor 370 and data transfer is completed after an appropriate time.

[0033] Drawing 4 shows the prefix header 400 generated by the primary storage controller 105. It is used by PDM104 in order to carry out the group division of the set of a self-description record after that. Each one, further, the journal of the oneself description record is carried out by between [SDM / 114] each time interval so that it may be given to secondary [DASD / 116] by the time sequence between each time interval. The prefix header 400 inserted in front of each group or a record set includes the actual primary-record set information sent to SDM114 to the sum total data length 401 and each record set for describing the sum total length of the prefix header. The operational research time stamp 402 is a time stamp showing the start time to the operational research set in which PDM104 is carrying out current processing. The operational research time stamp 402 is generated by PDM (following a SHISUPU REXX clock)104, when carrying out the "record set read" function to the set of the primary storage controller

105. The I/O time amount of writing primary [DASD / 106] is characteristic for the "record set read" of the each first storage controller 105. The operational research time stamp 402 is common over all storage controllers.

[0034] The time interval group number 403 is supplied by PDM104 in order to identify the time interval (it is ******** with a boundary by the operational research time stamp 402 and the record read time 407) to which a current record set (set of the record over all the primary storage controllers 105 to a given time interval group) belongs. The sequence number 404 in a group is discernment of a series of applications for the primary storage controller 105 to each record set in the given time interval group number 403 "I/O writing" of hardware supply having been carried out (as opposed to PDM104). Primary [SSID / 405] (sub storage discernment) identifies the specific primary storage controller of the primary storage controllers 405 to each record set peculiar. The secondary target volume 406 is assigned by PDM104 or SDM114 according to a performance consideration matter. The record read time 407 supplies the operational research time stamp common to all the primary storage controllers 105, and expresses the end time over the record set of a current interval.

[0035] Beforehand, system diagnostic information was collected based on an operator producing generalized trace facility (GTF). The GTF can collect system-hardware information based on the hook designed by the system. However, the trigger of the GTF is not carried out at the time of an error, data are caught according to an after-treatment function, therefore key debugging data are lost for selection impossible. When a ** platform offers the solution of hardware / microcode / software association, it sees from the complexity of a system increasing, and debugging technique needs to be improved. The solution which has

primary [of a large number like an equal remote copy (PPRC) and an extended remote copy (XRC)] and a secondary-storage controller needs distributed debugging for effective product quality maintenance. Such a distributed diagnosis combines hardware, a microcode, and software trace information effectively ranging over a system component, when a failure is detected (combination diagnosis).

[0036] The combination diagnosis is produced by new command word (CCW) like for example, diagnostic-state storage (DIAGNOSTIC STATE SAVE). The trigger of it is carried out by software or hardware at the time of early problem detection. Generating the "diagnostic-state storage" CCW makes a system catch critical software required to debug the problem, hardware, or a microcode control structure.

[0037] Then, such control structure information is combined with the continuous diagnostic trace information of that by the component tracing facility, and catches the perfect picture which combined the microcode / hardware environment, and the host software environment of an internal-storage controller ranging over one or more storage controllers. The external signs made to fail in specific I/O operation before this combination diagnostic solution carried out off-road one of the sense data about that specific hardware error to that hardware. In remote copy application, such an off-load is fully behind produced rather than the time of an actual error arising.

[0038] Drawing 5 shows the primary site 500. there, some applications 502 and 503 run in a host processor 501, and write renewal of a record in DASD (or many storage controllers 508 and 518 -- minding) 512 through the storage controller 508. The data mover 504 reads the prefix header relevant to the renewal of a record, and it, and in order to maintain a copy, it sends it to a secondary site. The secondary site (as [show / in drawing 1 / for example,]) is not shown in drawing 5 , in order to give

explanation easy. Such a secondary site is combined with the primary site 500 in asynchronous through a long distance in local (for example, enterprise system connectability (ESCON)) or through radiotelephony type transmission. The secondary-data mover 514 may exist in a host processor 501 like illustration, and may exist in a secondary site (not shown) depending on the requirements for a system.

[0039] The data mover 504 generates the "diagnostic-state storage" CCW in the case of detection of an error situation. In that case, a host processor 501 and the storage controller 508 collect ** software / hardware / microcode control structure information. TIF set to "OFF" while the instantiation-error situation had the truck information field (TIF) of a record count which are not equal to actual record count, and a record count equal to zero -- no, the inequality in - data, the TIF operation field showing the combination condition of an invalid, and the record length shown in TIF and the no false count in an end of record are included. Trace 507 exists in the data mover 504, and includes the separate address tooth space for memorizing a data mover trace diagnosis. The address tooth space in this example is a wrap around debugging tooth space. These data mover trace diagnoses include the return code which identifies the error type which made trace start, a flag condition, and header information.

[0040] Control 506 is also formed in the data mover 504. It is a data manager, it memorizes the information for controlling I/O operation, and control 506 carries out the map of the buffer tooth space to the storage controller of relation, and in order to form the adjustment group of renewal of a record for final transmission to a secondary site, it sorts I/O operation according to a time stamp. A buffer 505 is managed by control 506 and memorizes actual data (renewal of a record) besides TIF.

[0041] Generally the storage controller 508 contains the cache 509

segmented for the various renewal sessions of a record. Since a nonvolatile storage (NVS) 510 produces DASD high-speed write-in operation partially, it gives the storage tooth space for renewal of a record. Trace 511 makes the storage controller 508 keep proper status information. Other condition storage will not be performed, when this command is received and condition storage is advancing in the storage controller 508. The data transmitted as a result of the "diagnostic-state storage" CCW contain the following. Namely, the display of condition storage: Condition storage, general-purpose condition storage, or remote copy condition storage;
Session ID value;
Sequence number of the read record set of relation;
System abend code;
Reason-code; and the time-of-day stamp (TOD) which were detected by host software;
[0042] Thus, the approach for collecting data required in order to complete the problem decision in a complicated situation is performed in a storage controller. Information required to have mistaken especially the data read in the storage controller for the extended remote copy, and debug may exist in the queue structure in a cache. The event which sometimes produces a unique condition arose in the past. The probability with data required in order to debug a unique condition is made into max by giving a data mover session identifier and a sequence number from TIF of the read record set which the software is committing.
[0043] A data storage function assigns 64 K bytes of data in a cache. It is further divided into 16 K bytes of four blocks. The queue structure in a cache occupies the 16-K byte segment to 32 pieces. Therefore, a storage controller does not give sufficient condition storage storage to memorize all segments. Therefore, the microcode will choose the

suitable cache segment using the sequence number. The segment N+1 which adjoins the segment N including the identified record set and a degree is given. Furthermore, the first segment in the list and the last segment (therefore, four segments) will be given.

[0044] Drawing 6 is the flow chart showing the process which generates the "diagnostic-state storage" CCW for combining hardware, software, and a microcode control structure. The process starts at step 601 and the application which runs in one or more host processors generates renewal of a record for one or more storage controllers in step 603. In step 605, each storage controller memorizes updating directed to it to the internal cache of that, or NVS, therefore updates the control information (header) relevant to it. Next, a data mover reads renewal of a record into the buffer of that in step 607. An error situation is detected by the data mover in step 609. Based on detection of an error situation, a data mover generates the "diagnostic-state storage" CCW. The environment which contains the storage controller and host processor relevant to the error situation in it and coincidence is frozen.

[0045] In step 613, a storage controller starts discharge, in order to catch fault information. For example, a storage controller catches a control structure, a cache directory entry, remote one and the data queue which copied and changed, and a ** microcode data area. Moreover, in step 615, a host processor starts discharge, in order to catch fault information. For example, a host processor catches a host's internal-control structure. The caught fault information is memorized by the trace buffer of a storage controller and a data mover, and can be correlated in step 617 according to time amount, an event, etc. This data is analyzed in order to determine the cause of an error situation.

[0046] Drawing 7 shows the storage 701 which can memorize a data structure, computer instruction, or an application program in the form

which a processor can interpret easily and in which computer read is possible. Therefore, the step of drawing 6 is memorizable on the storage 701 in which computer read is possible as a series of computer instruction it can be directed to a storage system that performs this invention. The storage 701 in which computer read is possible is shown by the field as a floppy disk known well. It does not mean that the storage 701 in which computer read is possible is limited to a floppy disk, and is only limited by the function to memorize data or an instruction with high dependability rather in the form in which computer read is possible. For example, the computer memory of other suitable equal formats contains dynamic random access memory (DRAM), an optical disk, a magnetic tape, DASD, read-only memory (ROM), etc.

[0047] The above-mentioned explanation was limited to the error situation generated in the primary site. However, this concept is equally applicable ranging over a first order system and a second-order system. For example, the data mover (the 1st and 2nd data movers may be in one host processor) in a primary site can generate the "diagnostic-state storage" CCW for the storage controller in a secondary site, therefore can correlate the fault information from there.

[0048] If it summarizes, the data storage system which had a recovery function at the time of a distributed component and disaster will improve error debugging by adjusting fault information for catching the fault information from a suitable system component, and error analysis, when an error is detected. Fault information combines software, hardware, and a microcode control structure. A data storage system includes the secondary site for receiving a copy of data for the restorative purpose at the time of the primary site for processing data, and disaster. Each storage controller contains cache memory, a control buffer, and a trace buffer including the storage controller of plurality [site / primary]. A

primary host processor runs the application which generates renewal of a record, and in order to write in temporarily the renewal of a record of as opposed to it for the I/O operation to each renewal of a record, it sends it to two or more of the storage controllers. Moreover, a primary host processor assembles those renewal of a record in the group of a self-description record set including the primary data mover for reading renewal of a record in two or more of the storage controllers. When receiving the error code showing the generated type of an error situation from one of the system components, a data mover generates "diagnostic-state storage" channel command word (CCW) towards the system component relevant to the error situation, makes the related system component interrupt processing of renewal of a record temporarily, and makes fault information collect. Fault information is correlated according to the "diagnostic-state storage" CCW.

[0049] A primary data mover contains two or more buffers for memorizing the header of the control sections for managing the renewal of a record read into the trace queue for memorizing the fault information which carried out primary data mover relation, and the primary data mover, those renewal of a record, and relation. In order to receive renewal of a record from two or more primary storage controllers, two or more primary direct access storage devices (DASD) are combined with the primary storage controller. In order to receive the group of a self-description record set, a secondary-data mover is combined, and the "diagnostic-state storage" CCW is answered. A secondary host processor is combined with a secondary-data mover, and it also answers the "diagnostic-state storage" CCW. Two or more secondary-storage controllers are combined with a secondary host processor, and the "diagnostic-state storage" CCW is answered. In order to memorize the group of a self-description record set, secondary

[two or more / DASD] is formed.

[0050] Probably this invention is shown in a detail with reference to the suitable example, and it will be clear to this contractor that it is possible to make various change in a format and a detail, without deviating from the pneuma and the technical range of this invention also although it explains and excels. For example, storage may be a tape or optical equipment. Therefore, it is possible to carry out without deviating from having taught modification of the example in the future of this invention in this application.

[0051] As a conclusion, the following matters are indicated about the configuration of this invention.

[0052] (1) Make it the storage system which has the dispersed system component containing a host processor. Said host processor runs the software application which generates renewal of a record. When it has a data mover, and was combined with the 1st storage controller and an error situation arises in said storage system, In the storage system which performs an approach effective in a machine in order that said data mover may adjust said dispersed problem decision between system components (a) The step which produces the I/O operation for the renewal of a record generated by said software application, (b) The step which memorizes said renewal of a record in said 1st storage controller according to said produced I/O operation, (c) The step which maintains the control information relevant to said renewal of a record in said 1st storage controller, (d) The step which reads into said data mover said control information related [which was related and record-updated] from said 1st storage controller in preparation for copying said renewal of a record remotely, (e) The step which detects the storage system error situation in said data mover, (f) The step which generates diagnostic-state storage channel command word (CCW) for said host

processor and said 1st storage controller from said data mover, (g) Said fault information in each system component is an approach characterized by correlating according to the detected error situation including the step which catches fault information in said host processor and said storage controller.

(2) Said step (f) is an approach given in the above (1) characterized by including the step which freezes system environment until said step (g) is completed.

(3) Said approach is an approach given in the above (2) characterized by including the step which analyzes fault information further in order to determine the cause of the (h) error.

(4) Said step (e) is an approach given in the above (2) characterized by including the step which receives the error code showing an error type in said data mover.

(5) Said step (g) is an approach given in the above (2) characterized by including the step which memorizes host-processor fault information in the circulation trace buffer in said data mover.

(6) Said step (g) is an approach given in the above (2) characterized by including the step which memorizes the 1st storage controller fault information in the circulation trace buffer in said 1st storage controller.

(7) Said fault information is an approach given in the above (2) characterized by including the hardware, the software, and the microcode control structure for giving debugging.

(8) Said step (f) is an approach given in the above (2) characterized by including the step which generates said diagnostic-state storage CCW for the 2nd storage controller in a remote location.

(9) Make it the storage in which the computer read for memorizing data mover application in order to answer a storage system error and to combine a condition storage diagnosis ranging over many system

components is possible. In the storage containing the host processor combined with the storage controller for said system component to manage the renewal of a record generated by the software application to a direct access storage device (DASD) The means for producing the I/O operation to the renewal of a record generated in said software application, The means for memorizing said renewal of a record for said storage controller according to said produced I/O operation, The means for maintaining the control information relevant to said renewal of a record in said storage controller, The means for reading into said data mover said control information related [which was related and record-updated] in preparation for copying said renewal of a record remotely, The detection means for detecting a storage system error situation and communicating said error situation to said data mover, The condition storage means for generating diagnostic-state storage channel command word (CCW) for said host processor and said storage controller from said data mover, The prehension means for catching fault information in said host processor and said storage controller, An implication and said fault information in each system component of the system components of said large number are data mover application characterized by correlating according to the detected error situation.

(10) Said condition storage means is data mover application given in the above (9) characterized by freezing the software application relevant to said error situation until the operation and said prehension means of said storage controller complete prehension of fault information.

(11) Said detection means is data mover application given in the above (9) characterized by receiving the error code showing the error type of said error situation in said data mover.

(12) Said prehension means is data mover application given in the above (9) characterized by memorizing host-processor fault information in the

circulation trace buffer in said data mover.

(13) Said prehension means is data mover application given in the above (12) characterized by memorizing storage controller fault information in the circulation trace buffer in said storage controller.

(14) Make it the data storage system for adjusting fault information between [relevant to the generated error situation] system components. In the data storage system containing one or more storage controllers combined with the nonvolatile storage for said system component to memorize renewal of a record The host processor which runs a software application is included. Said software application is what generates renewal of a record and sends the I/O operation for the usual storage in said nonvolatile storage to said one or more storage controllers. Said host processor reads said renewal of a record in said one or more storage controllers. It is a thing containing the data mover for assembling said renewal of a record in a group, in order to send to the remote storage system for recovery at the time of disaster. Said data mover is what receives the error code showing the type of the produced error situation from one of said the system components. It is what generates the condition storage command for making said related system component interrupt processing of renewal of a record temporarily in order that said data mover may collect fault information in said related system component. And that said fault information is what is correlated between system components according to a condition storage command and said data mover The trace queue for memorizing the fault information relevant to said data mover, The data storage system characterized by including the control section for managing the renewal of a record read into said data mover, and two or more buffers for memorizing the header to which said renewal of a record and they relate.

(15) A data storage system given in the above (14) characterized by including the remote data mover combined so that the group of renewal of a record might be received from said data mover.
(16) Said nonvolatile storage is a data storage system given in the above (14) characterized by being a direct access storage device (DASD).
(17) It is a data storage system given in the above (15) characterized by answering said condition storage command when an error situation produces said remote data storage system in said remote data storage system including the remote data storage system combined so that the group of renewal of a record might be received from said remote data mover for storage.
(18) Said remote data storage system is a data storage system given in the above (17) characterized by including a remote host processor, the remote storage controller combined with said remote host processor, and the remote nonvolatile storage for being combined with said remote storage controller and memorizing the group of said renewal of a record.
(19) Said fault information is software, hardware, and a data storage system given in the above (14) characterized by including a microcode control structure.
(20) The secondary site which receives a copy of data for recovery at the time of the primary site for processing data, and disaster, It has software, hardware, and the distributed-system component that had the microcode control structure incorporated. It is made a recovery data storage system at the time of the disaster which gives error debugging by carrying out the trigger of adjustment and prehension of the fault information relevant to the generated specific error situation distributed-system component mutual [said]. Two or more storage controllers by which each has cache memory, a control buffer, and a trace buffer, It is what sends the I/O operation to each renewal of a

record to said two or more storage controllers in order that said application may generate renewal of a record and may write said renewal of a record in two or more storage controllers including the host processor which runs application. Said host processor is a thing containing the primary data mover for reading said renewal of a record in said two or more storage controllers, and assembling it in the group of a self-description record set. Said primary data mover is what receives the error code showing the type of the produced error situation from one of said the system components. In order that said primary data mover may collect fault information It is what generates the diagnostic-state storage channel command word for making said related system component interrupt processing of renewal of a record temporarily towards the distributed-system component relevant to said error situation. Said fault information is what is correlated according to diagnostic-state storage channel command word, Two or more primary direct access storage devices combined with said two or more primary storage controllers (DASD), The secondary-data mover which is connected so that the group of a self-description record set may be received, and answers said diagnostic-state storage channel command word, The secondary host processor which is memorized by said secondary-data mover and answers said diagnostic-state storage channel command word, Two or more secondary-storage controllers which are combined with said host processor and answer said diagnostic-state storage channel command word, Secondary [for memorizing the group of said self-description record set / two or more / DASD] is included. Said primary data mover The trace queue for memorizing the fault information relevant to said primary data mover, It is a recovery data storage system at the time of the disaster characterized by including the control section for managing the renewal

of a record read into said primary data mover, and two or more buffers for memorizing the header to which said renewal of a record and they relate.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] It is the block diagram showing a remote dual copy system including a primary site and a secondary site.

[Drawing 2] It is the block diagram of the storage controller in the remote dual copy system shown in drawing 1 .

[Drawing 3] It is the block diagram of the storage pass processor contained in the storage controller shown in drawing 2 .

[Drawing 4] It is drawing showing the header-data structure for the renewal of a record in the storage controller shown in drawing 2 .

[Drawing 5] It is the block diagram of the primary site storage system by the suitable example of this invention.

[Drawing 6] It is the flow chart which shows how to adjust the fault information between system components of drawing 5 according to the suitable example of this invention.

[Drawing 7] It is the memory in which computer reading for memorizing the computer instruction which performs the approach of drawing 6 is possible.